

# ABSTRACT

## Pre-Processing Influences on Multivariate Data Analysis of Time-Of-Flight Secondary Ion Mass Spectroscopy Data and Images

Thérèse Gerri-Ann Lee

Time-of-Flight Secondary Ion Mass Spectroscopy (ToF-SIMS) has been used in the investigation and examination of biological compounds for more than three decades. Currently, ToF-SIMS data analyses are quite daunting tasks and there exists problems with their interpretation. Multivariate Statistical Analyses (MVA), such as Principal Components Analysis (PCA) and Maximum Auto-correlation Factors (MAF) have proven to be valuable aids for interpreting ToF-SIMS data but the results are sensitive to a variety of pre-processing steps. Pre-processing techniques include calibrating the mass scale, dead time correction, peak selection, binning, normalisation, and scaling of the data. The main aim of this research was to explore some of the pre-processing techniques which can be applied to ToF-SIMS data and their impact on the multivariate data analysis of data.

Although many pre-processing techniques are available for data analysis, this research explored the techniques which have not been extensively researched inclusive of binning, peak selection and dead time correction. Subsequent to the pre-processing of the data, it was further processed using PCA and MAF. Besides investigating processing techniques, the topography on a surface was investigated. The data used in this research included synthetically generated data and actual data obtained from biological samples and all data analyses were performed in MATLAB®. From the research conducted, it was found that the pre-processing method selected for the analysis of ToF-SIMS data and the manner in which it was performed highly influenced the results from the multivariate data analyses.

Keywords: Thérèse Gerri-Ann Lee; Multivariate data analysis; pre-processing; ToF-SIMS; PCA; MAF; binning; peak selection; topography.